

ASME 2014 Verification & Validation Symposium

Sandia Verification & Validation
Challenge Workshop

A Bayesian Inference Approach to the Validation Challenge Problem



Wei Chen¹ Daniel W. Apley² Wei Li¹ **Zhen Jiang**¹ Shishi Chen¹
Professor Professor Visiting Scholar PhD Candidate Visiting Scholar

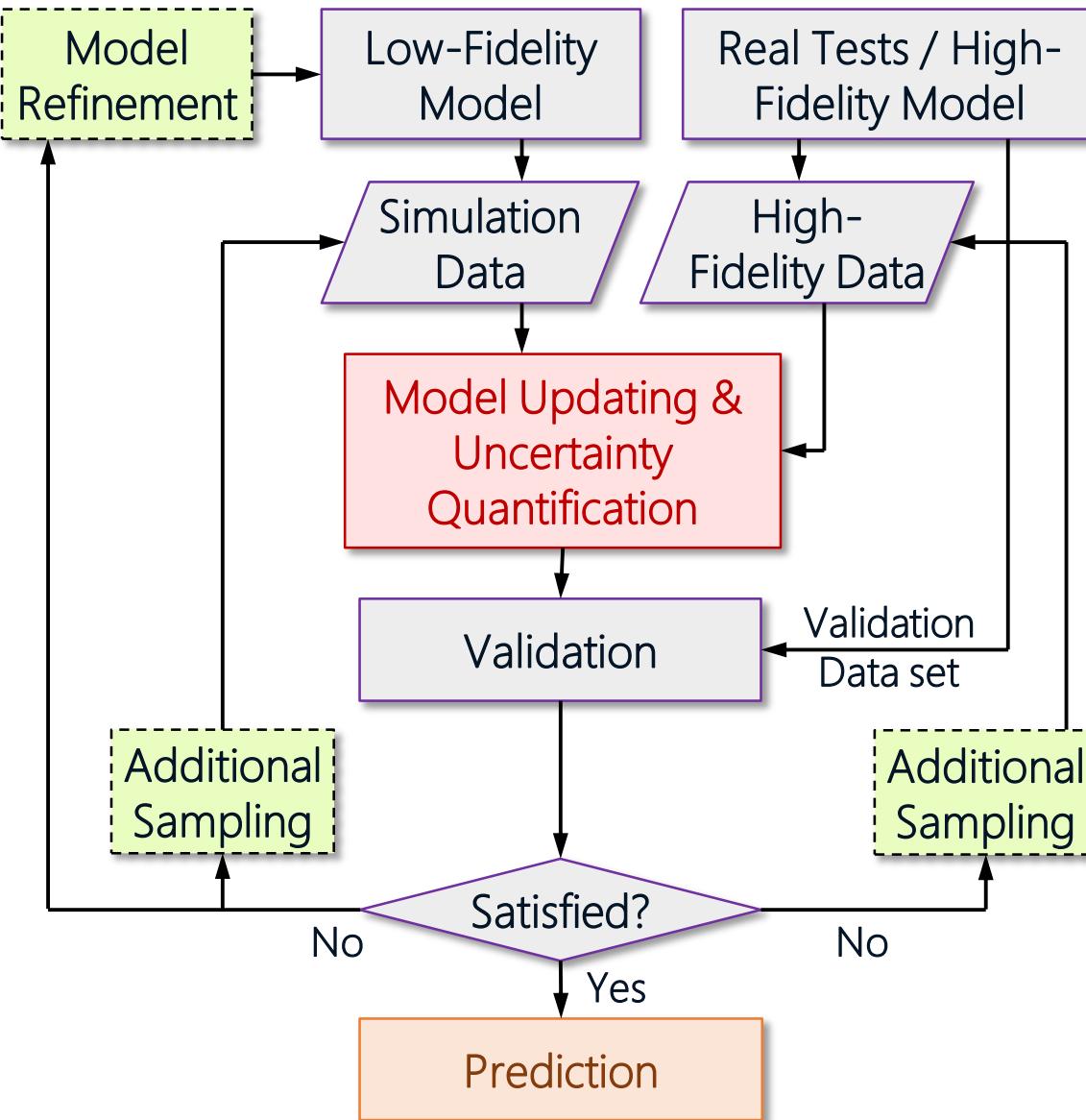
¹Department of Mechanical Engineering, Northwestern University, Evanston, IL

²Department of Industrial Engineering & Management Sciences, Northwestern University, Evanston, IL

Introduction

> General Framework for Model Updating and Uncertainty Quantification

Introduction

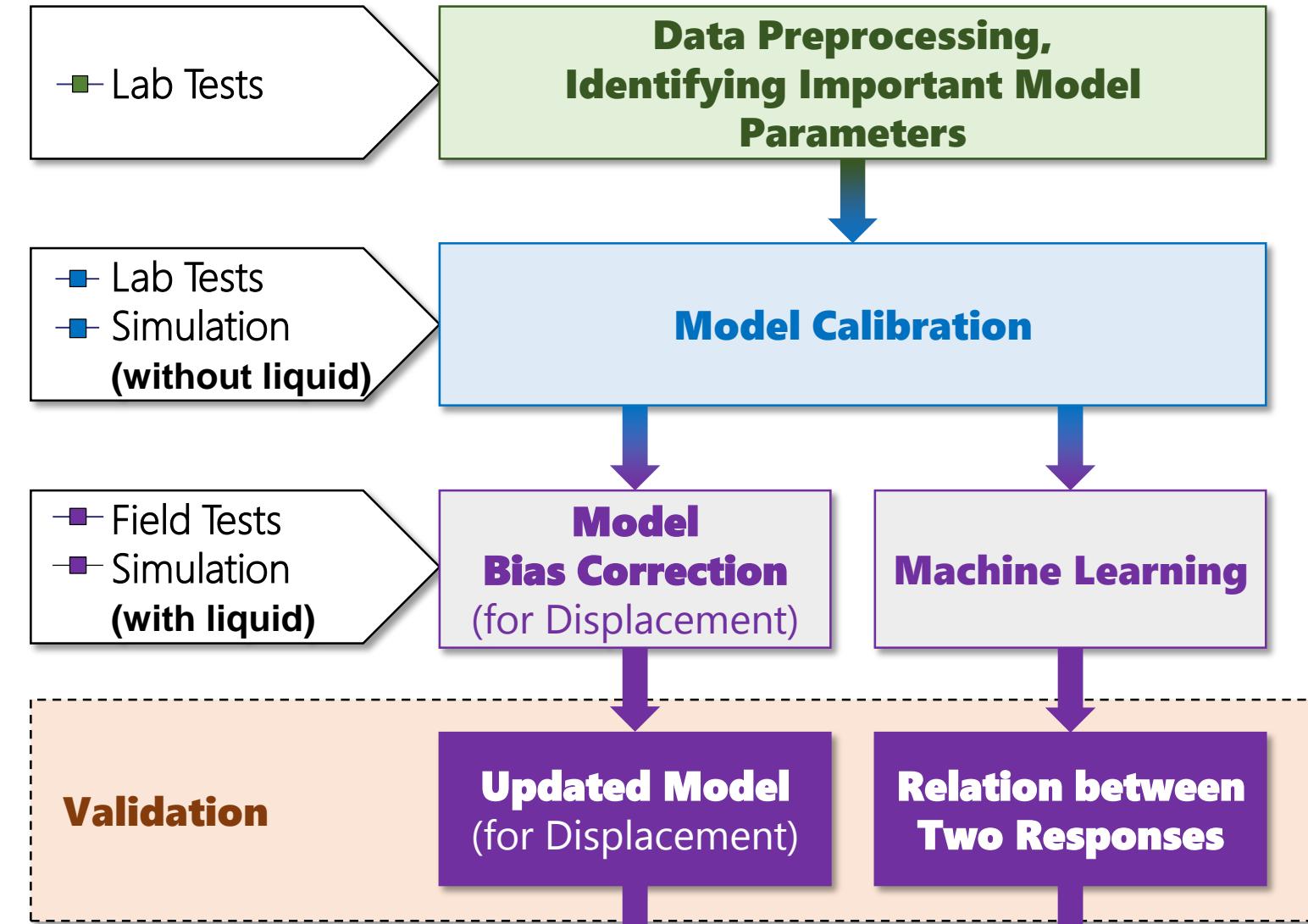


> Challenges

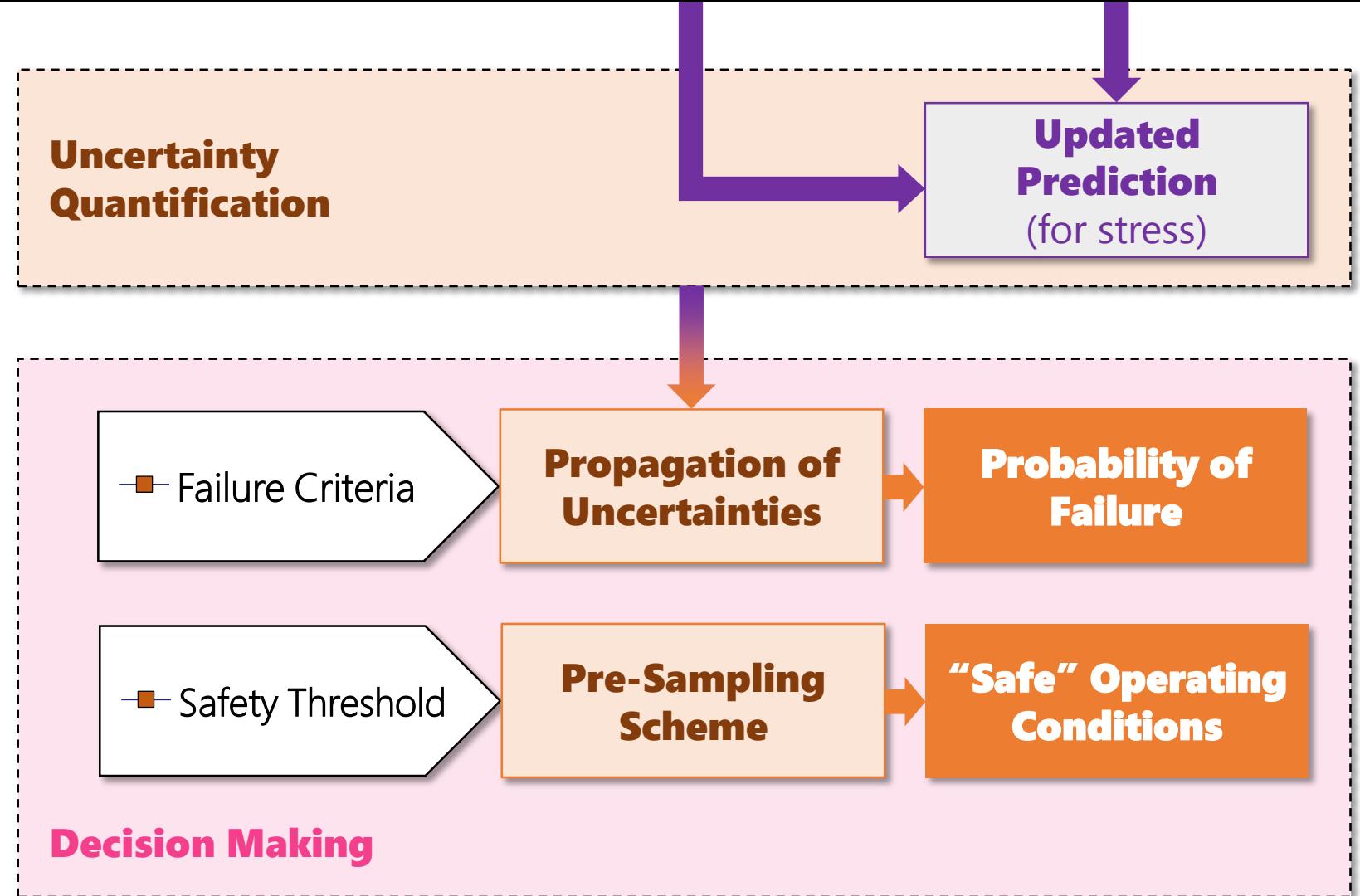
- Unknown parameters
- Abundant (Redundant) data
- Imperfect simulation model
- No experimental data of von Mises stress
- No expert opinions
- Uncertainty quantification
- Uncertainty propagation

Challenges

- ─ Unknown parameters
- ─ Abundant (Redundant) data
- ─ Imperfect simulation model
- ─ No experimental data of von Mises stress
- ─ No expert opinions
- ─ Uncertainty quantification
- ─ Uncertainty propagation



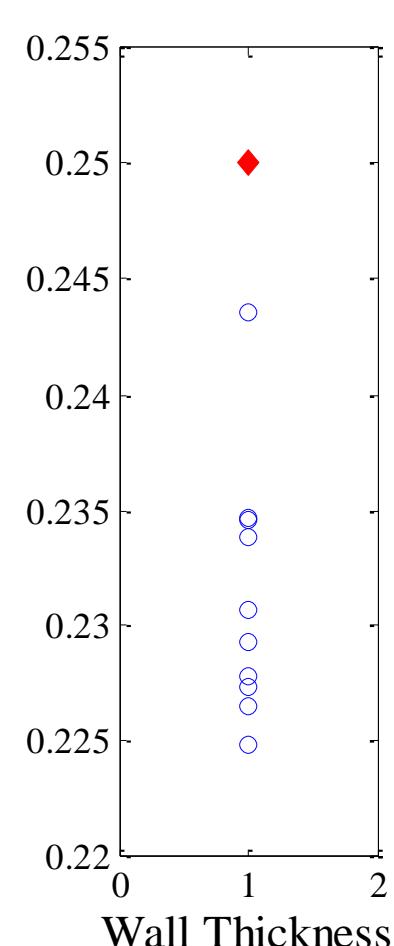
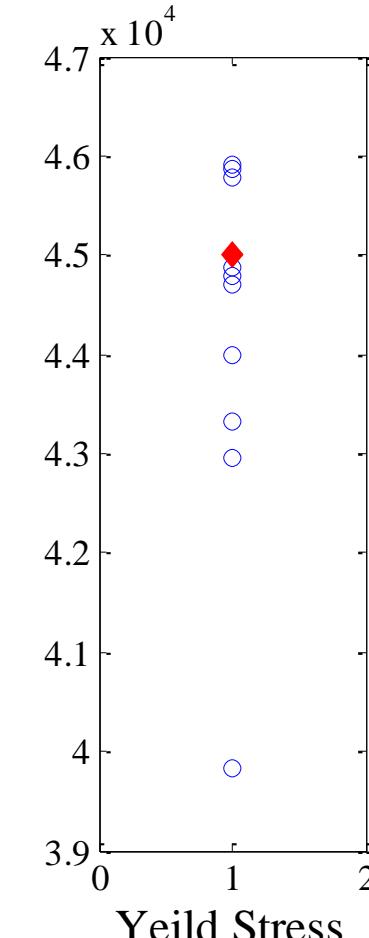
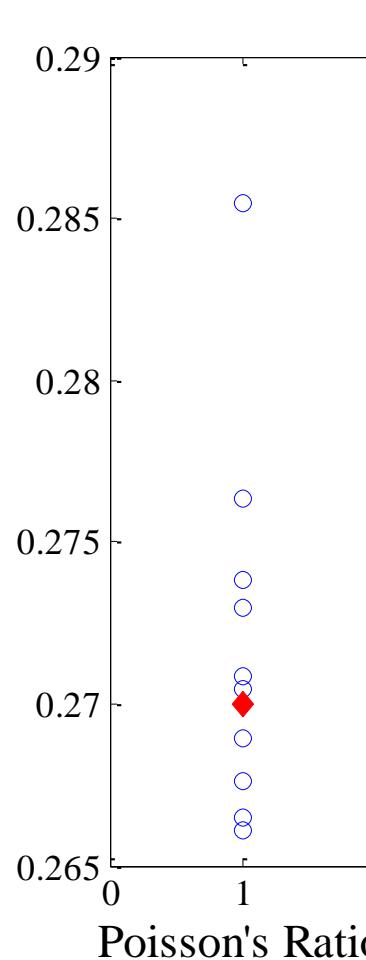
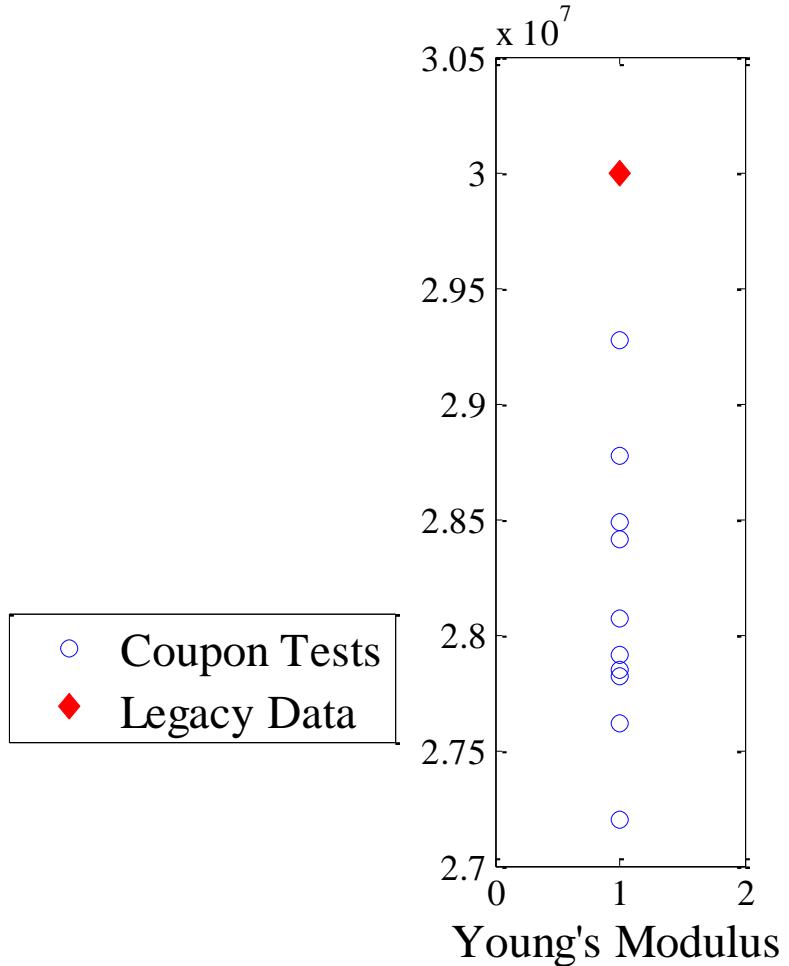
Introduction



Data Preprocessing and Sensitivity Analysis

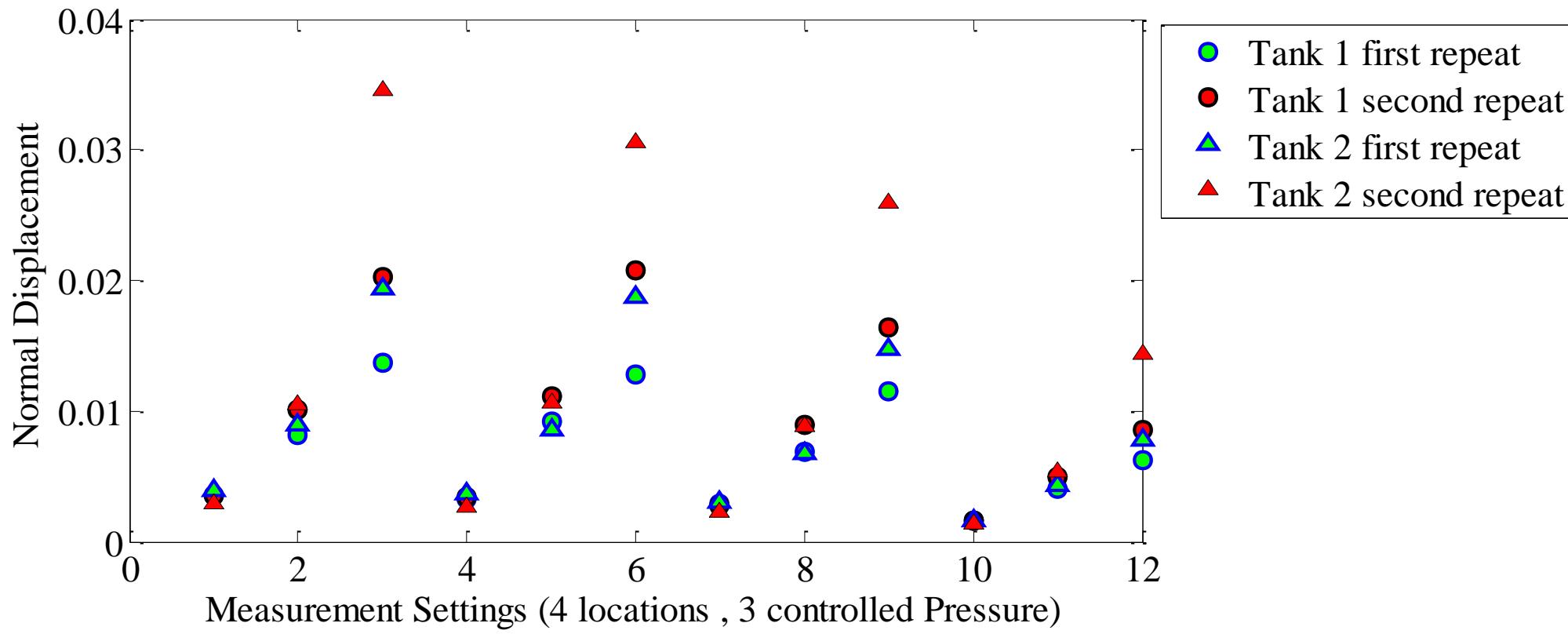
> Data Associated with Model Parameters

- Comparison between Legacy Data and Coupon Tests
- Conclusion: Legacy data will not be used.

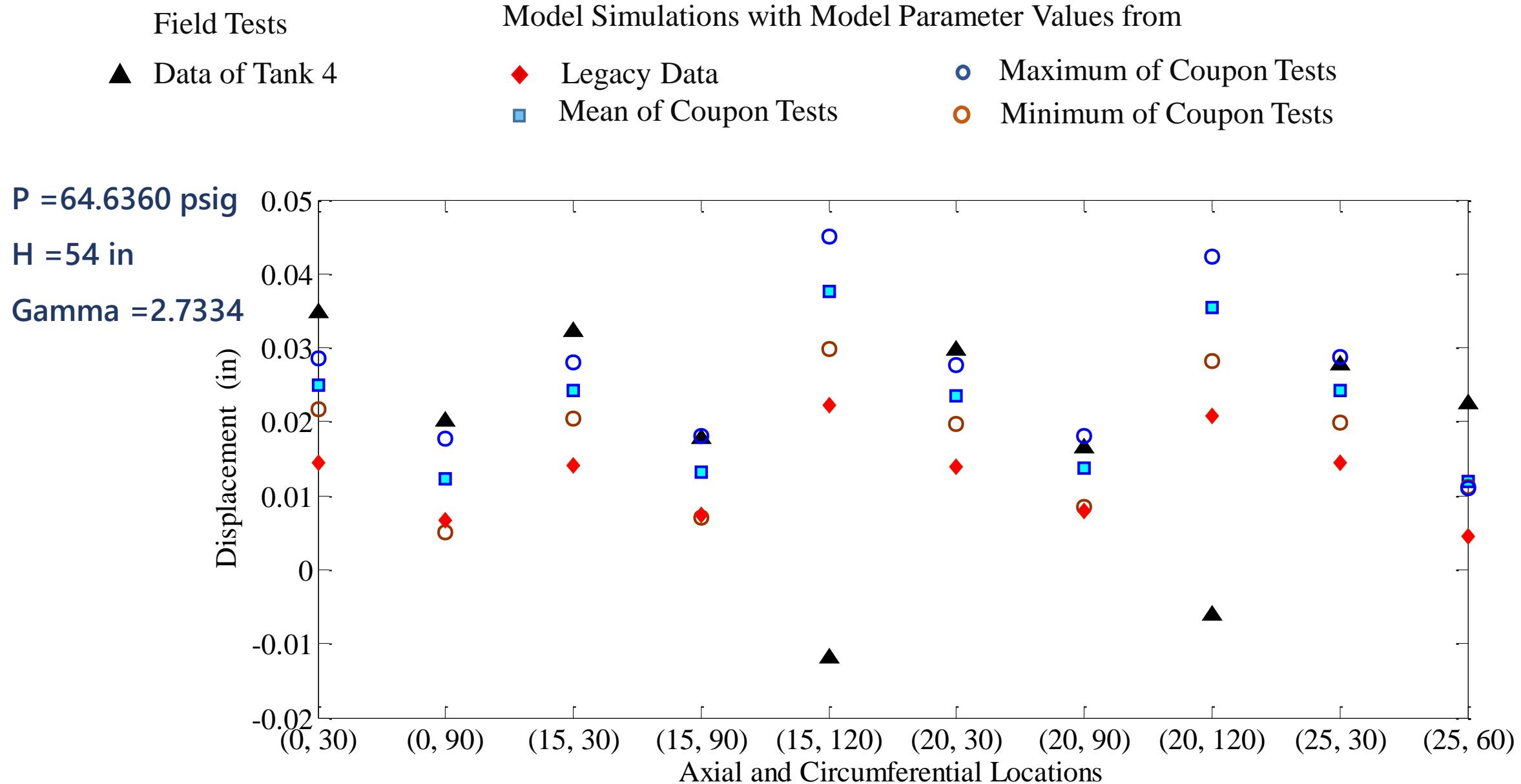


> Reduce Experimental Uncertainty by Data Preprocessing

- Outliers in measurements should be removed.



> Why Model Calibration is Needed?



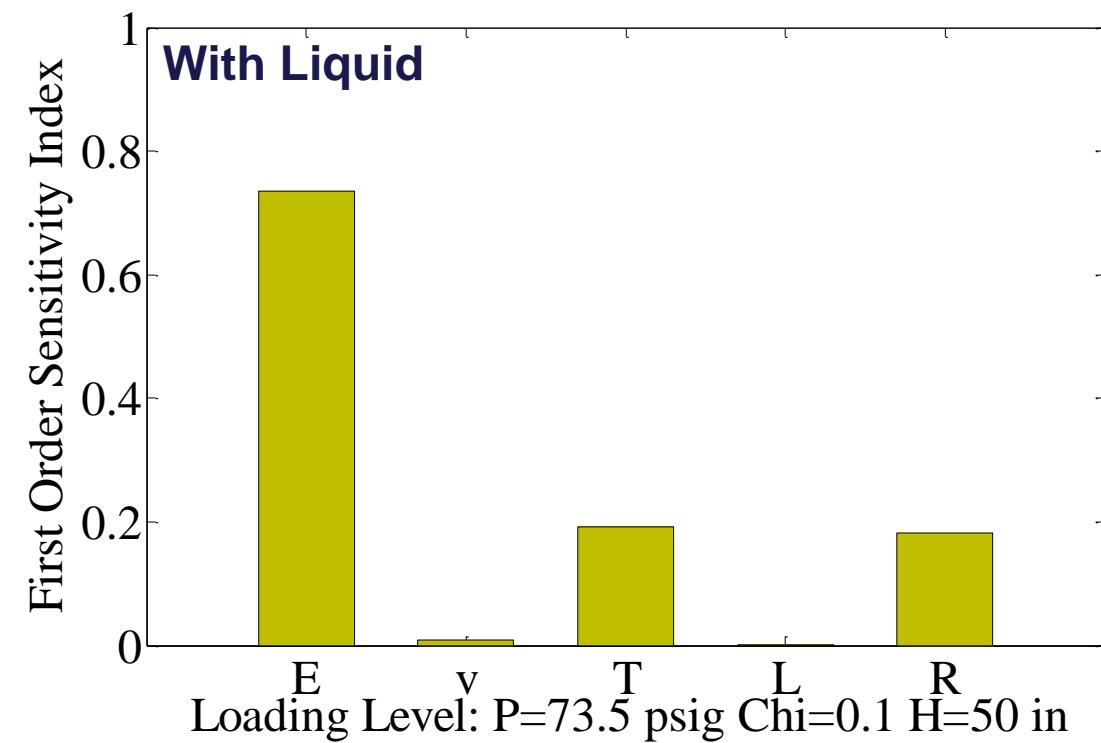
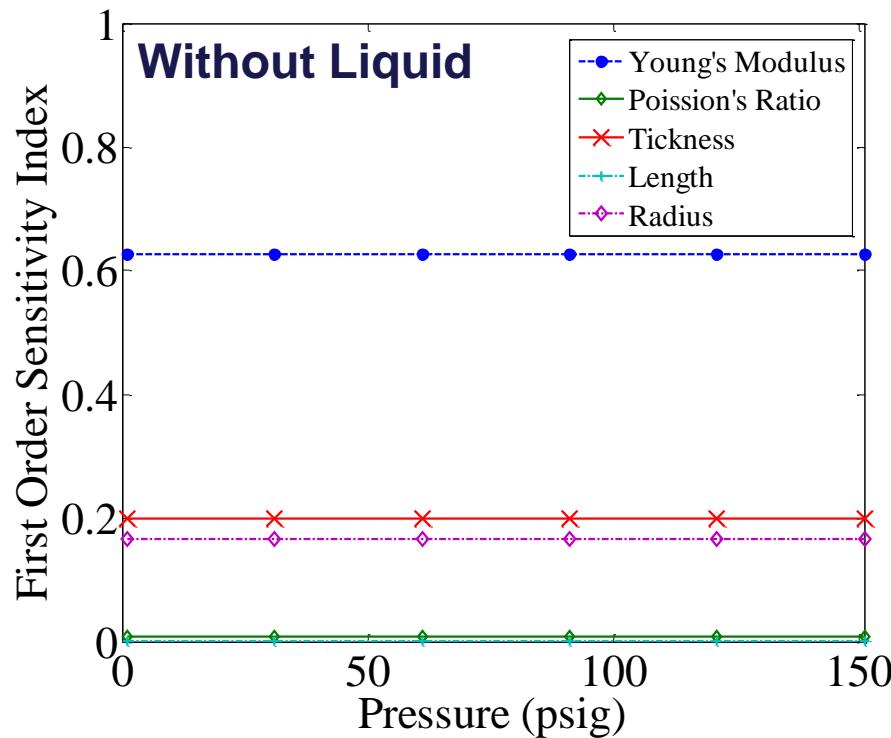
> Global Sensitivity Analysis

▪ Simulation Model

$$[w, \sigma] = M(x, \varphi, P, \gamma, H, E, v, L, R, T, m)$$

▪ Main Sensitivity Index

$$S_i = \frac{V_{X_i} (E_{\mathbf{X}_{-i}} (Y | X_i))}{V(Y)}$$

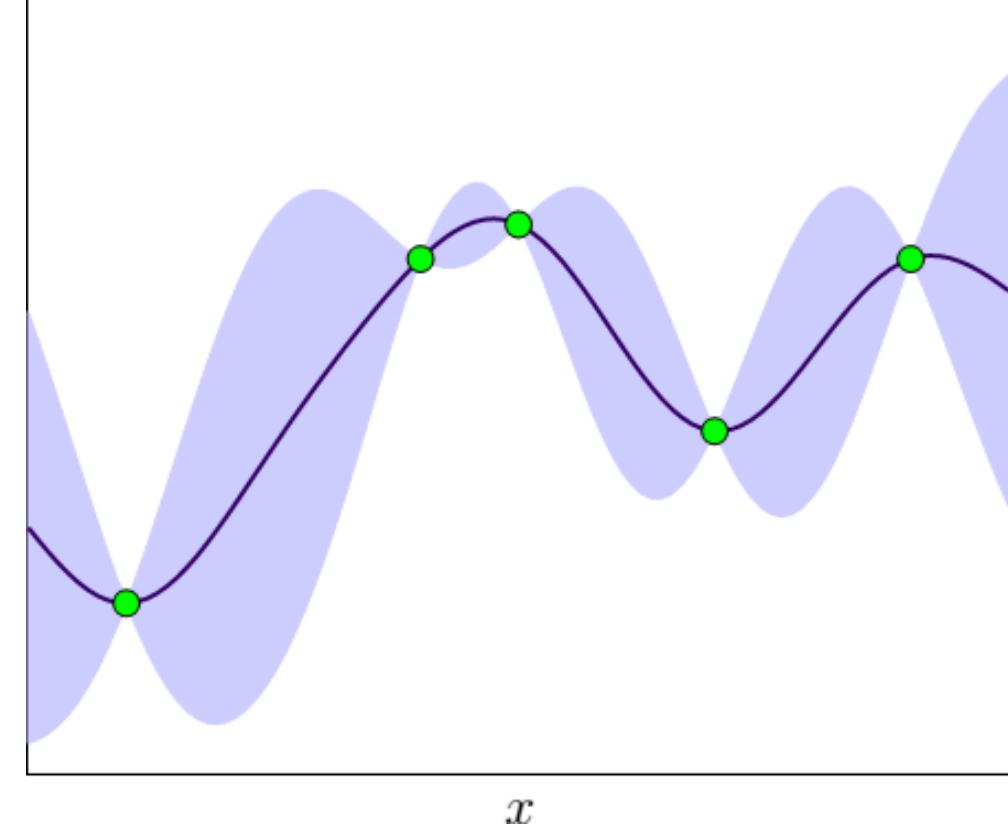
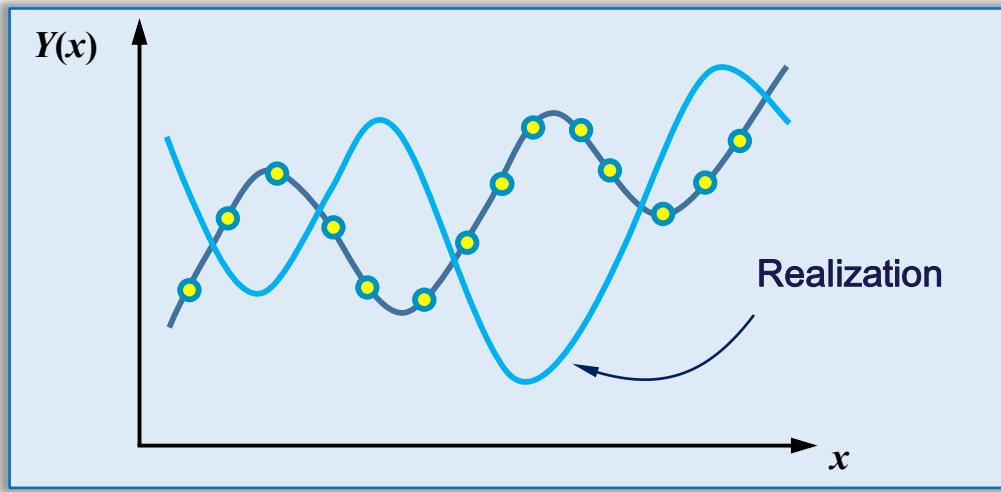


- Conclusion: E and T are most important parameters that need calibration.

Spatial Random Process Based Model Calibration, Bias Correction, and Validation

> Spatial Random Process (SRP) Modeling

SPATIAL RANDOM PROCESS (SRP)



EXAMPLE: GAUSSIAN PROCESS (GP)

$$\bullet \begin{bmatrix} Y(\mathbf{x}_1) \\ \vdots \\ Y(\mathbf{x}_n) \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} m(\mathbf{x}_1) \\ \vdots \\ m(\mathbf{x}_n) \end{bmatrix}, \begin{bmatrix} V(\mathbf{x}_1, \mathbf{x}_1) & \cdots & V(\mathbf{x}_1, \mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ V(\mathbf{x}_n, \mathbf{x}_1) & \cdots & V(\mathbf{x}_n, \mathbf{x}_n) \end{bmatrix} \right)$$

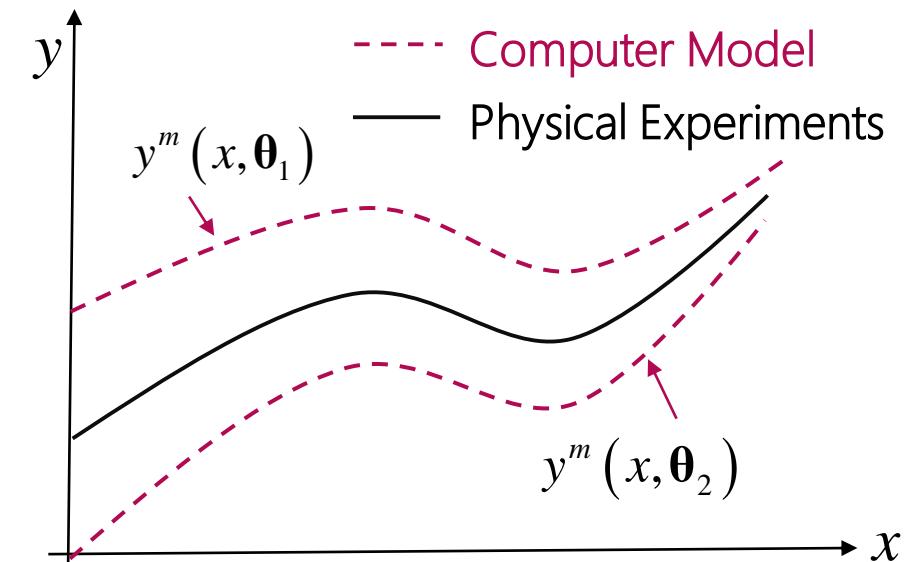
- Widely used for metamodeling (surrogate modeling) and regression

> Model Calibration and Modular Bayesian Approach

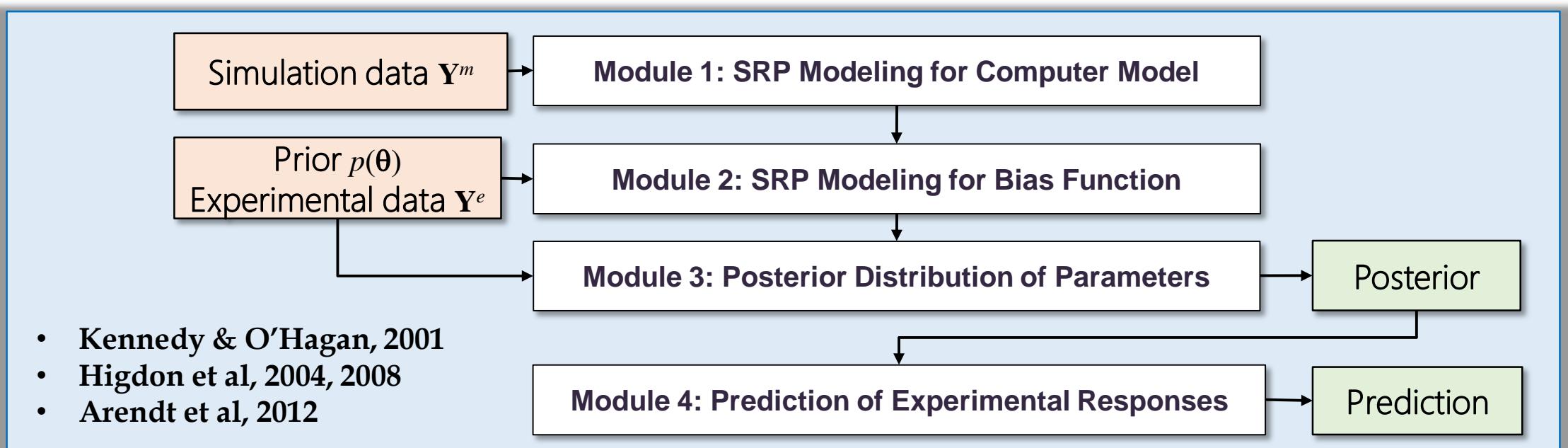
FORMULATION

$$y^e(\mathbf{x}) = \underbrace{y^m(\mathbf{x}, \boldsymbol{\theta}^*)}_{\text{Computer Model}} + \underbrace{\delta(\mathbf{x})}_{\text{Experimental Error}} + \underbrace{\varepsilon}_{\text{Bias Function}}$$

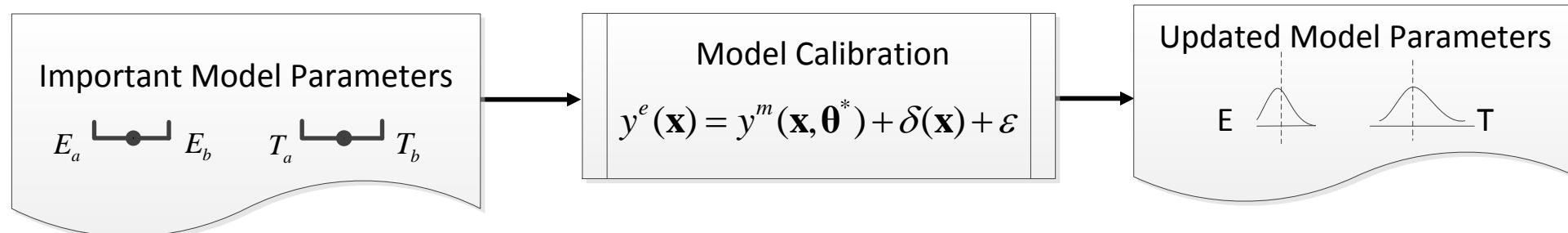
Input Variable Computer Model Experimental Error
 Physical Experiments Calibration Parameters Bias Function



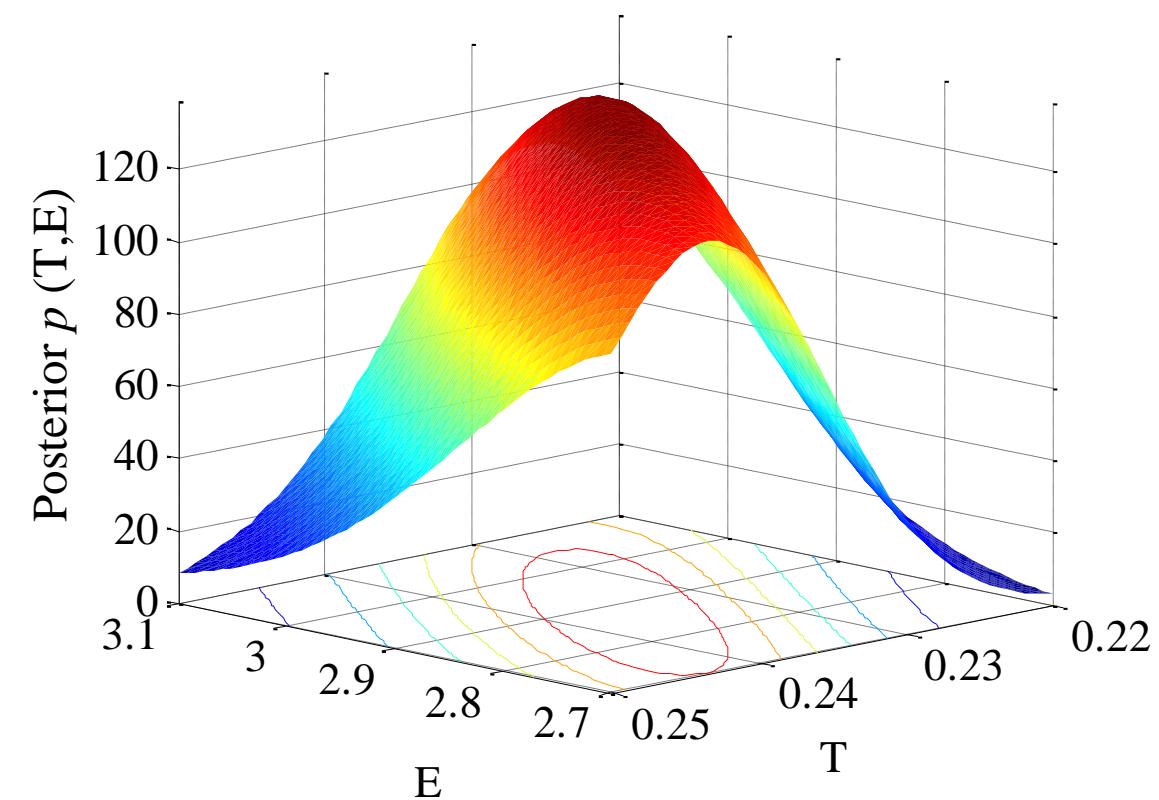
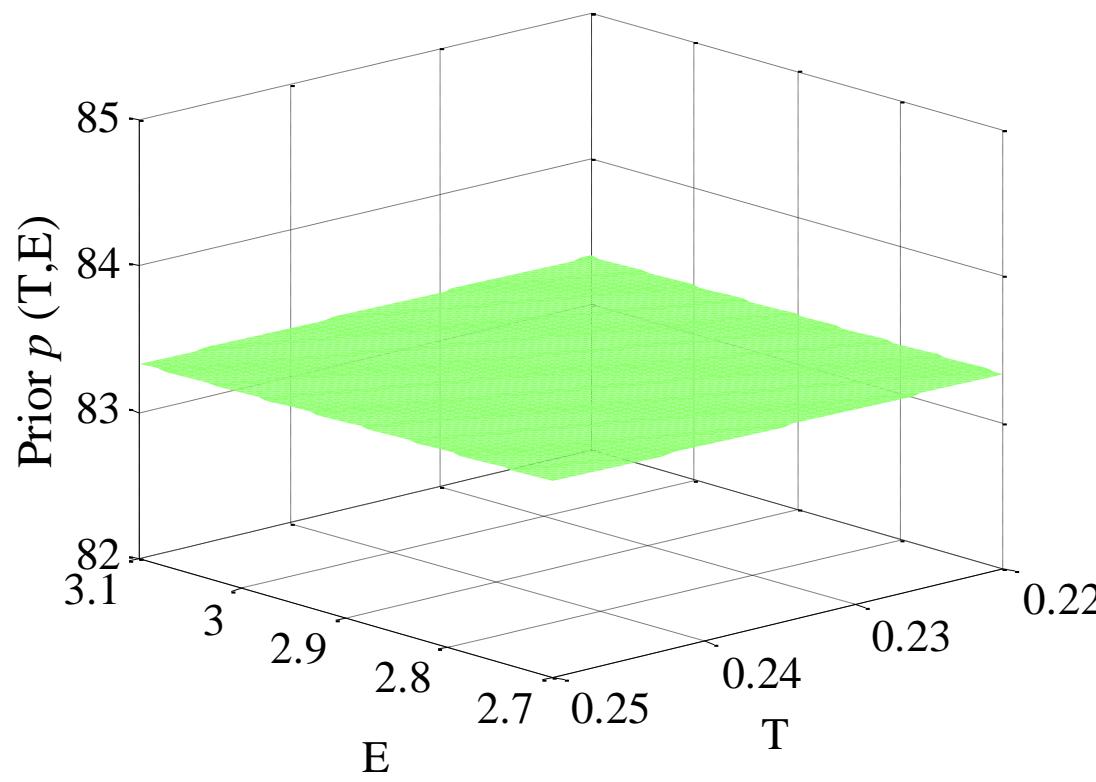
MODULAR BAYESIAN APPROACH



Model Calibration Using Model Simulations without Liquid

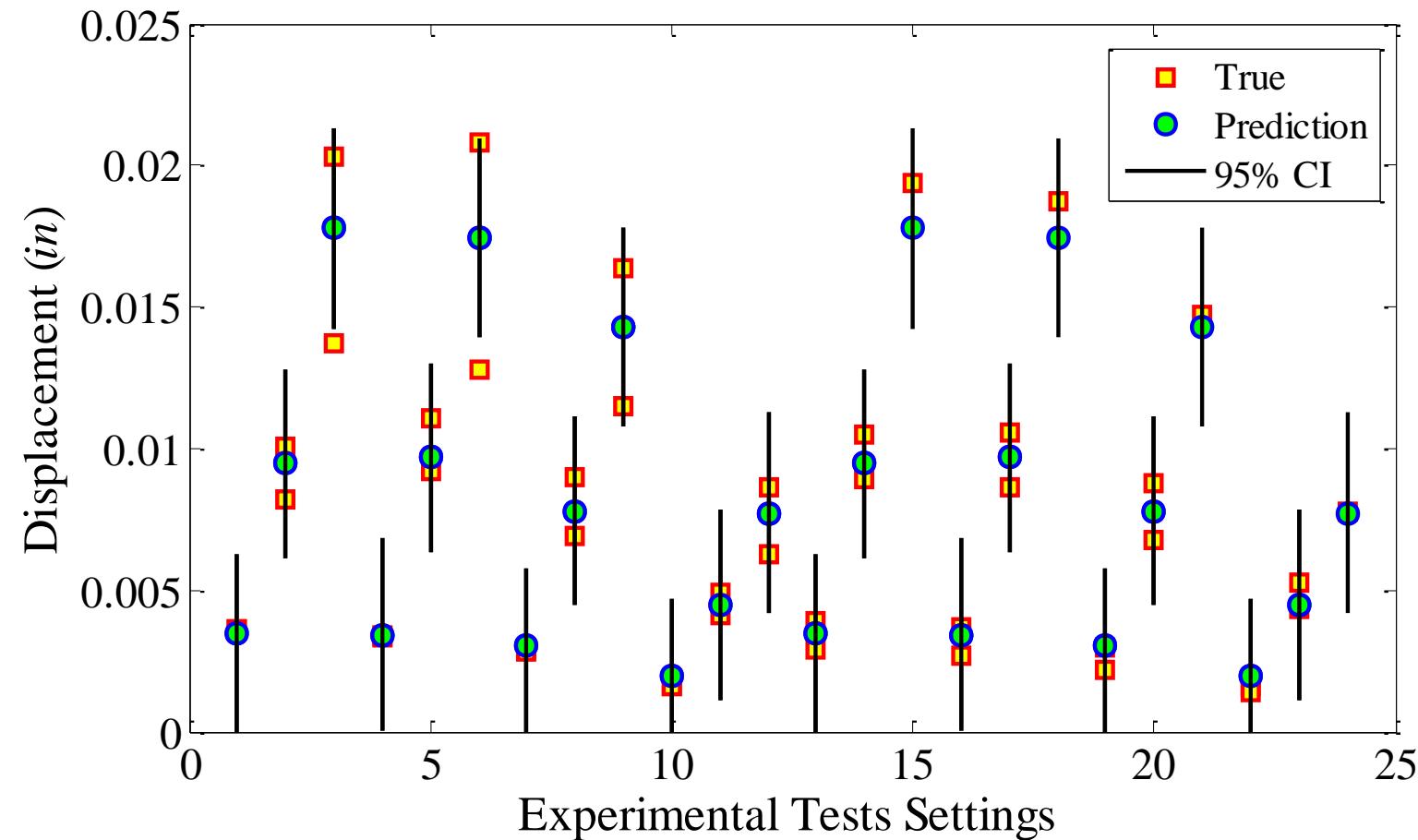


$T : [0.225, 0.235]$ $E : [2.720, 2.927]$

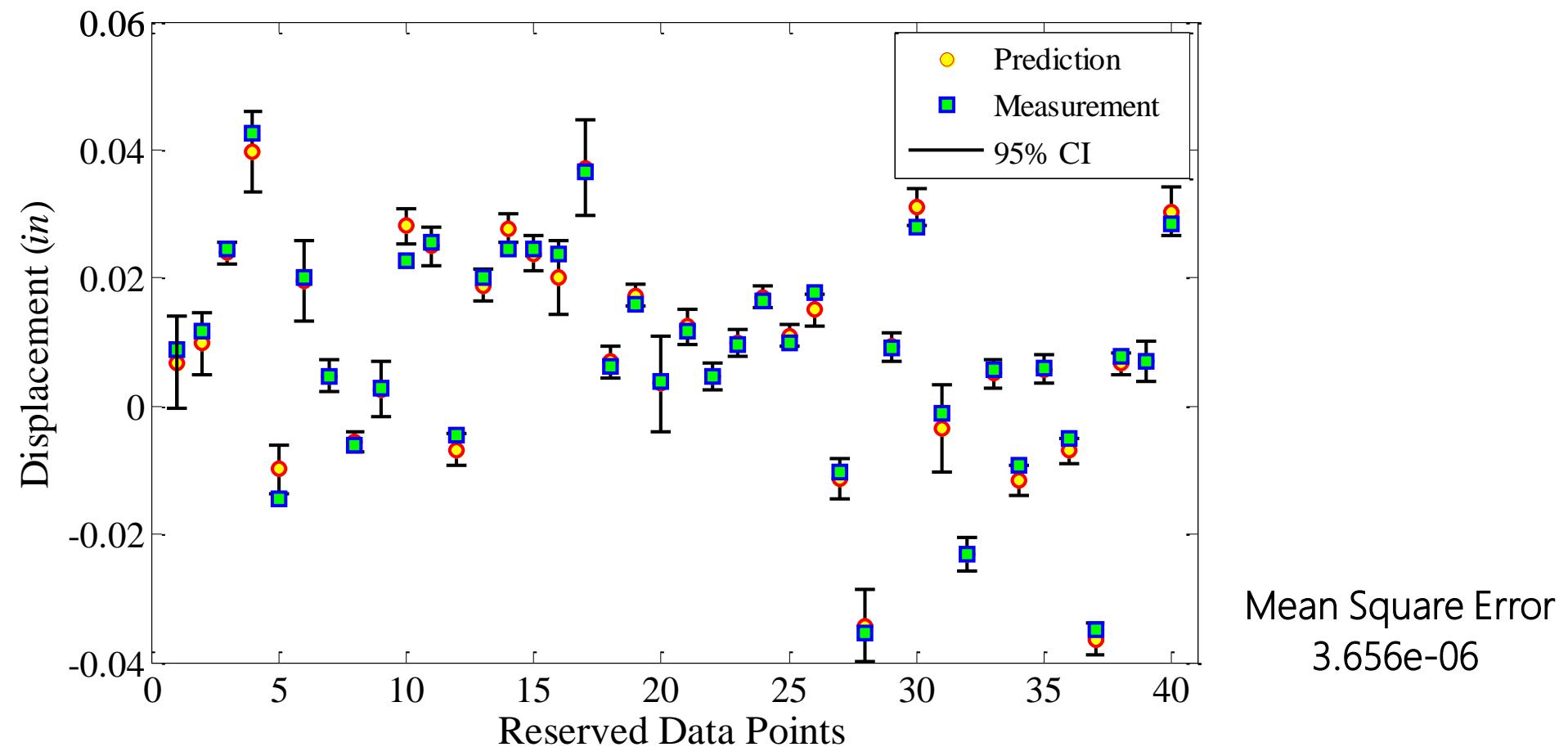
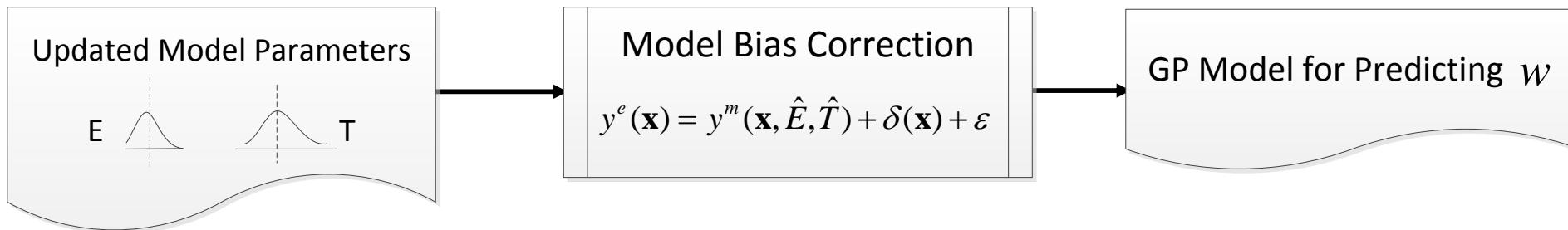


Validation of Model Calibration Result

Calibration Parameters	Estimation	Mean	Standard Deviation	Coefficient of Variation
Young's Modulus (psi)	2.8796×10^7	2.9001×10^7	0.11×10^7	0.038
Wall Thickness (in)	0.2359	0.2353	0.007577	0.032



Bias Correction Using Model Simulations with Liquid

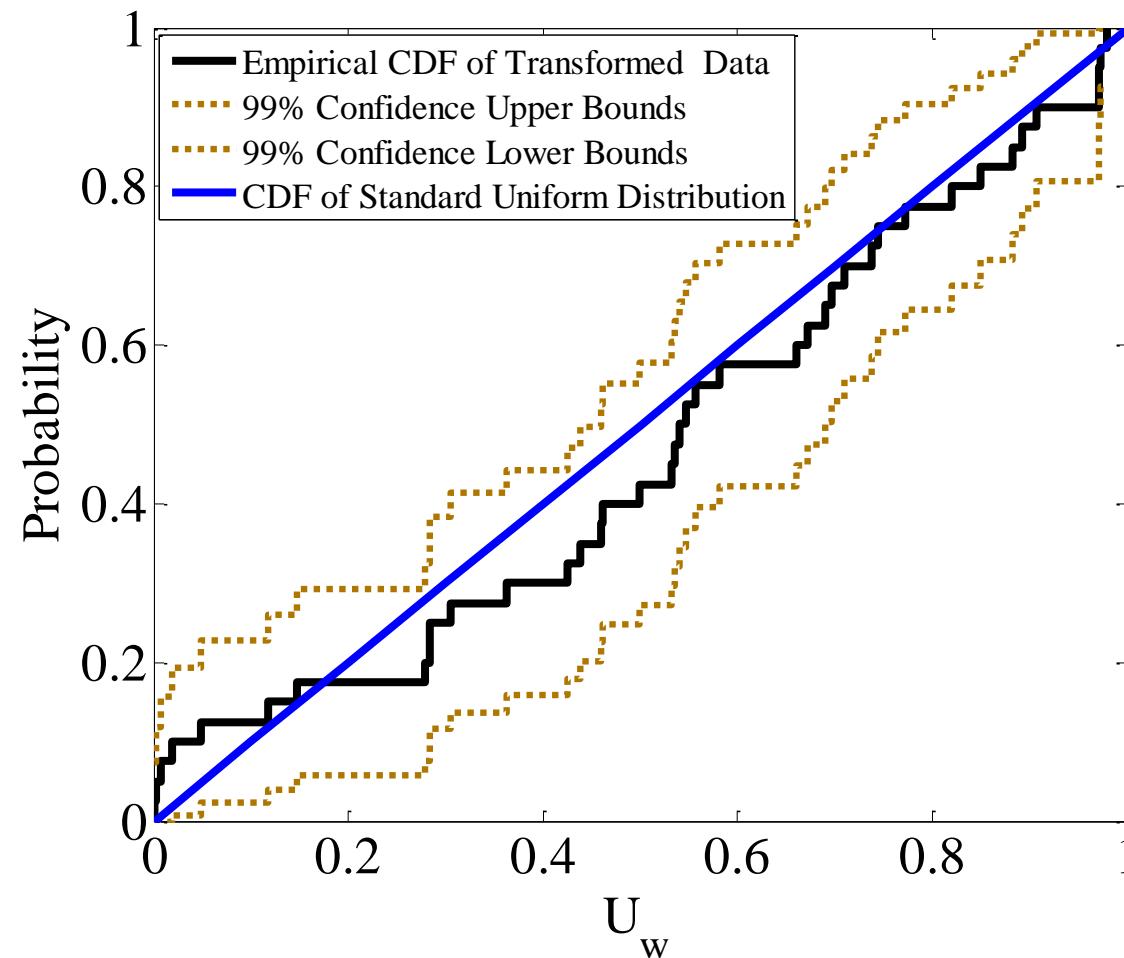




Validation of Model Bias Correction Result

U-POOLING METRIC

- Pools original data and predictions from multiple input settings into a single metric
- Metric value: 0.0356 (smaller-the-better)

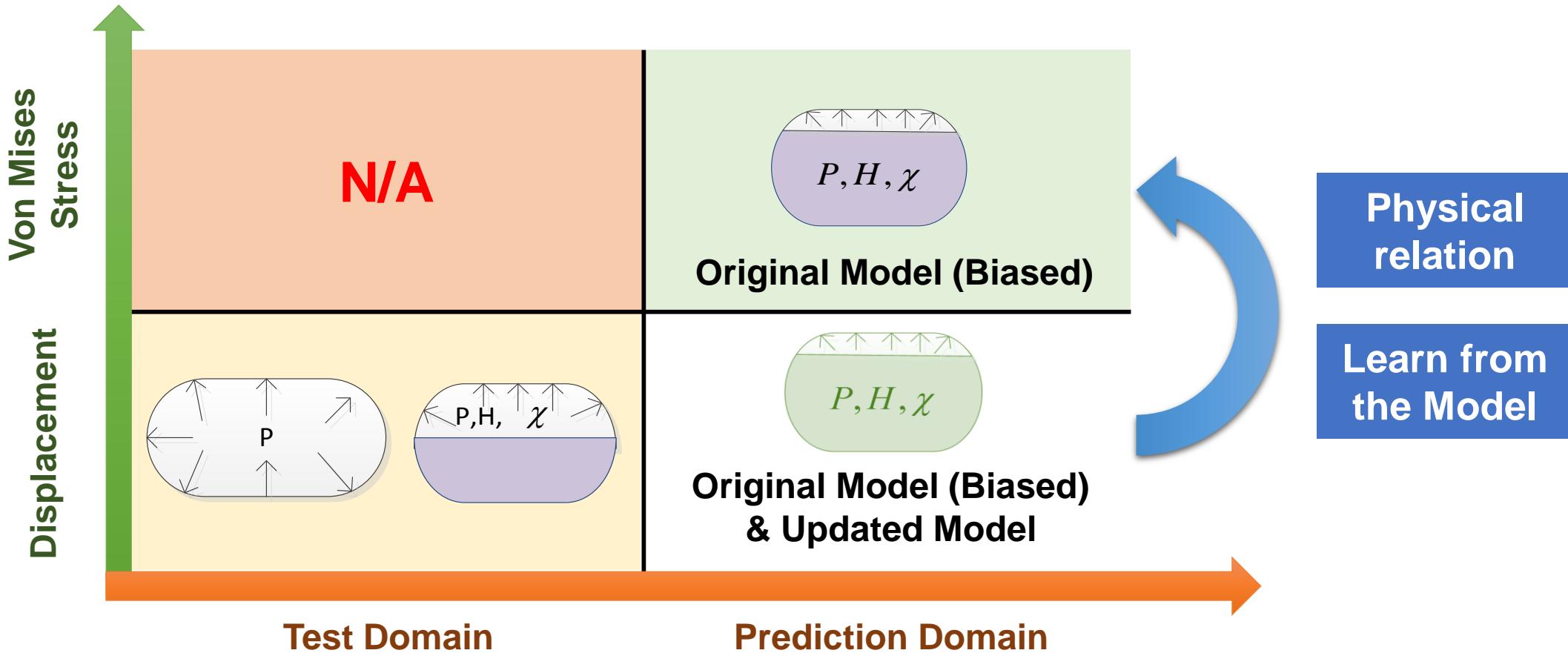


Machine Learning for the Relation Between Two Responses



Relationship Between Two Responses

EXPERIMENTS & PREDICTIONS



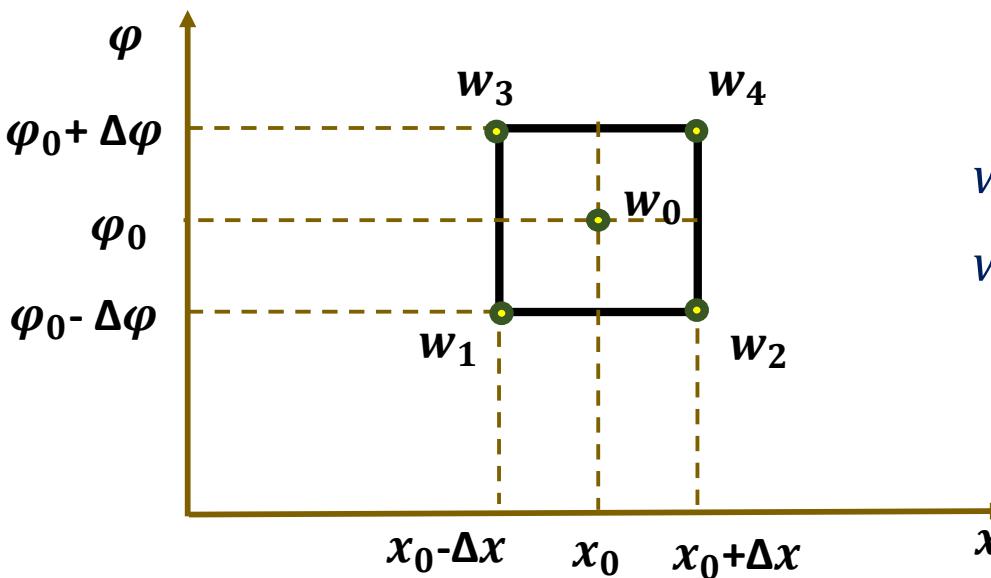
> Machine Learning from Simulations

SIMULATION MODEL

- Biased, but built upon sophisticated physical theories
- Assumption: The relation between w and σ in the model is correct

LEARNING STRESS AS A FUNCTION OF DISPLACEMENT

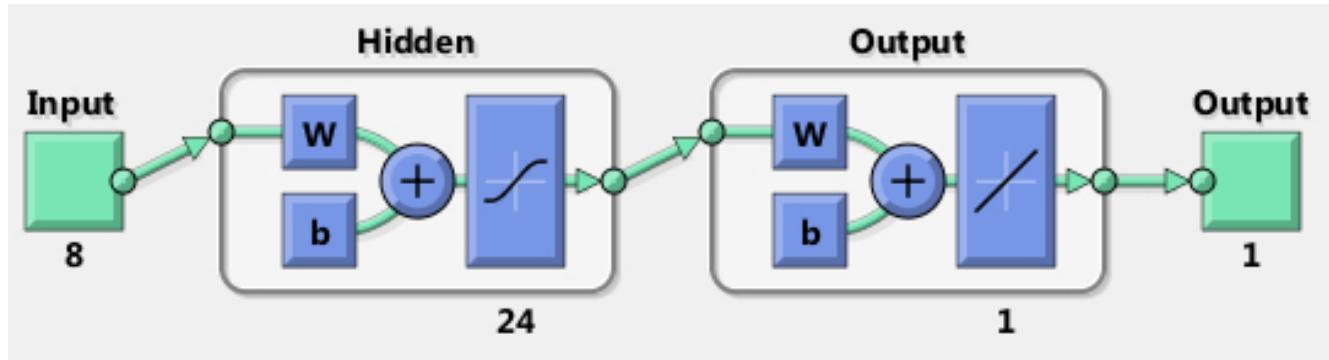
- $\sigma = f_0(x, \varphi, H, w_0, w_1, w_2, w_3, w_4)$



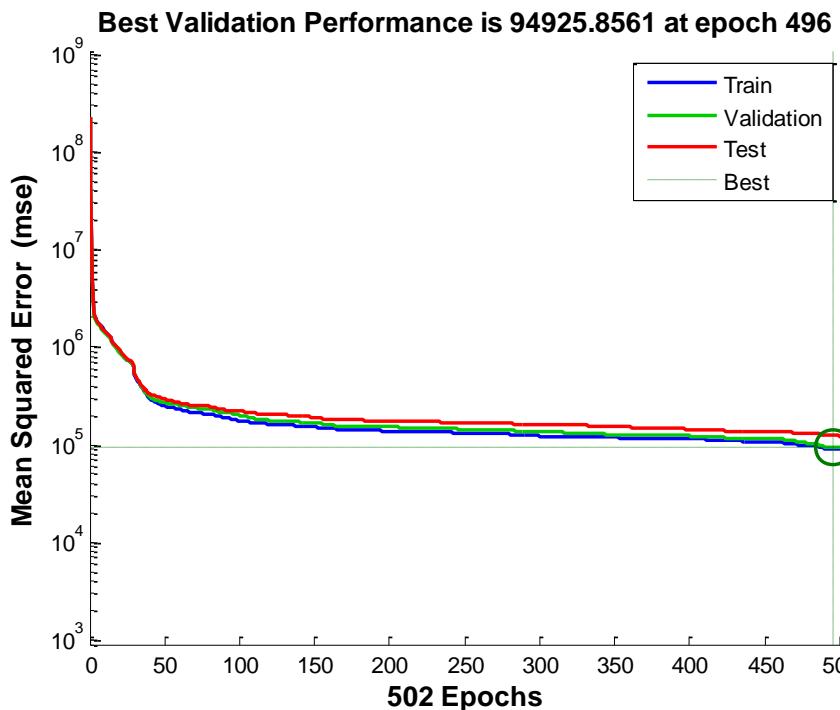
w_0 : Displacement at x , and φ
 $w_1 \sim w_4$: Displacement at locations near x , and φ

> Neural Network and Regression Performance

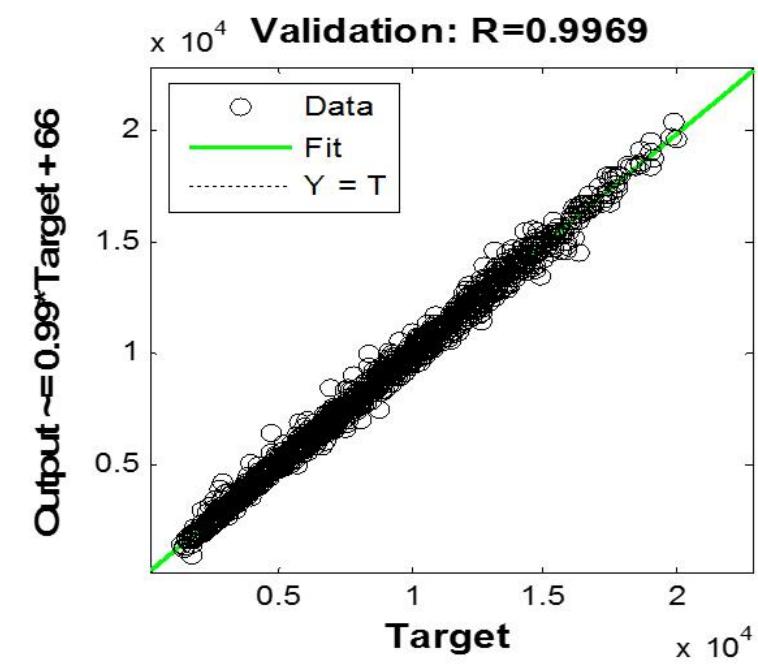
NEURAL NETWORK



PERFORMANCE: MEAN SQUARE ERROR



VALIDATION: CORRELATION COEFFICIENT



Uncertainty Quantification, Uncertainty Propagation, and Predictions

> Multiple Sources of Uncertainty

Sensitivity Analysis

Data Pre-processing
and Sensitivity Analysis

Experimental Variability
Parameter Uncertainty

Model Calibration

Updated Model Parameters


Parameter Uncertainty

Bias Correction

GP Model for Predicting
 w

Experimental Variability
Interpolation Uncertainty
Model Discrepancy

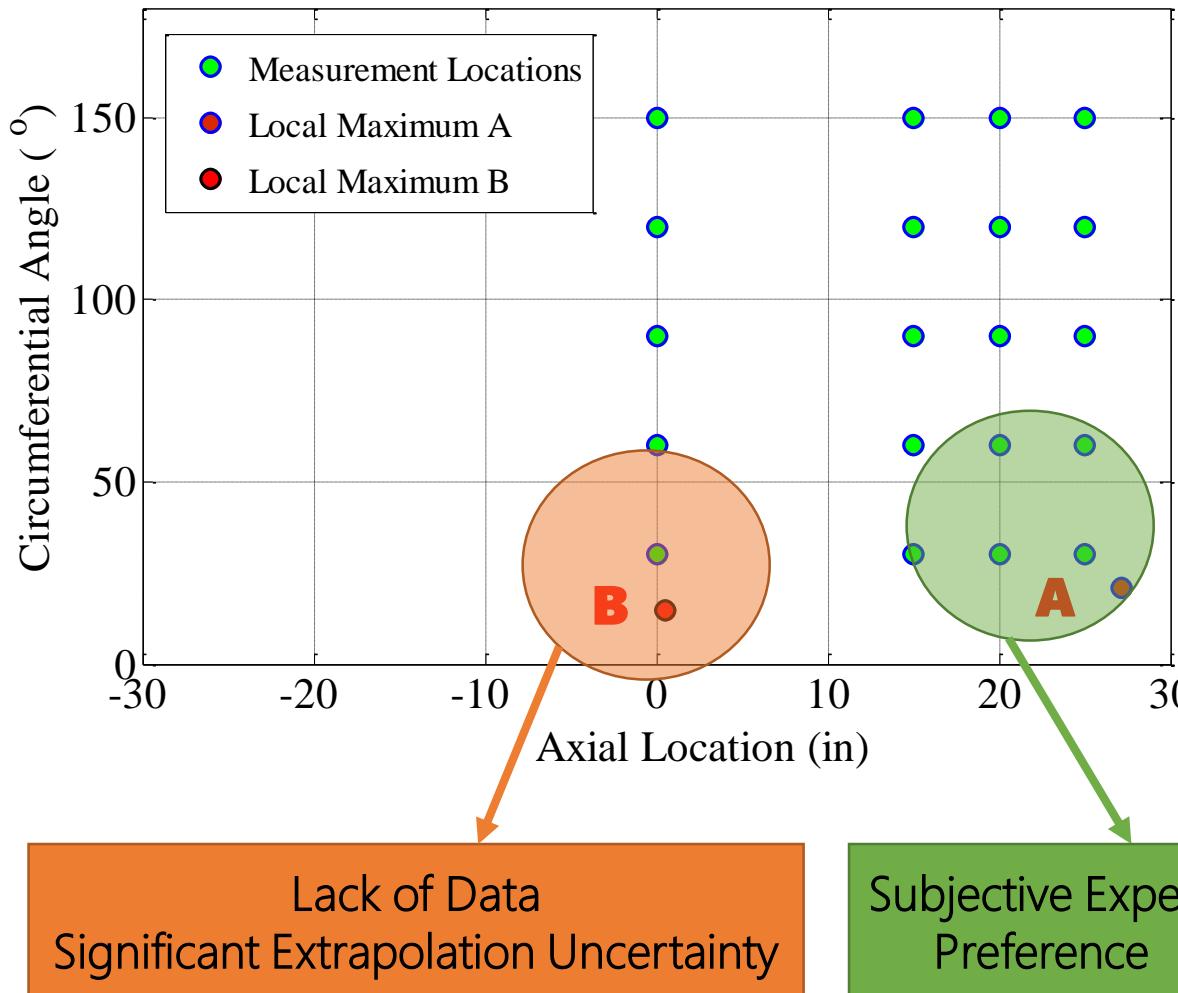
Machine Learning

Function Relation between
 w and σ

Numerical Uncertainty

> Maximize Stress & Probability of Failure at Nominal Condition

OPTIMAL SOLUTIONS



Local Maxima	x and ϕ Location	Mean
A	(27.116, 10.801)	2.367×10^4
B	(0.554, 14.752)	2.4257×10^4
Standard Deviation		$P(\text{Failure})$
A	8.182×10^3	0.006038
B	2.1875×10^4	0.180927

Conclusion: Tank will fail at location A with failure probability 0.6%.

> Safe Operating Conditions and Pre-Sampling Scheme

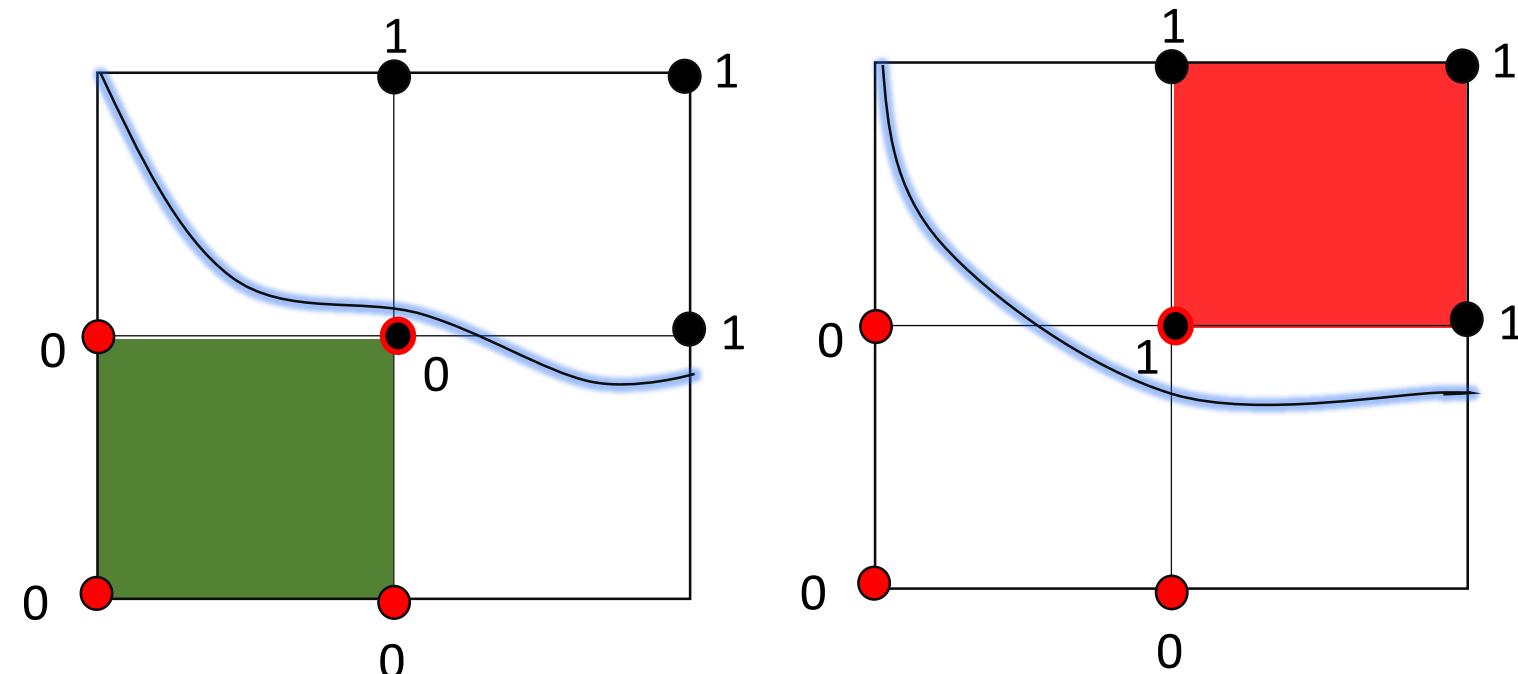
OPERATING CONDITIONS

- In terms of P , H and χ
- P , H and γ are variables associated with loading levels,
the larger the loading levels, the higher $P(\text{Failure})$



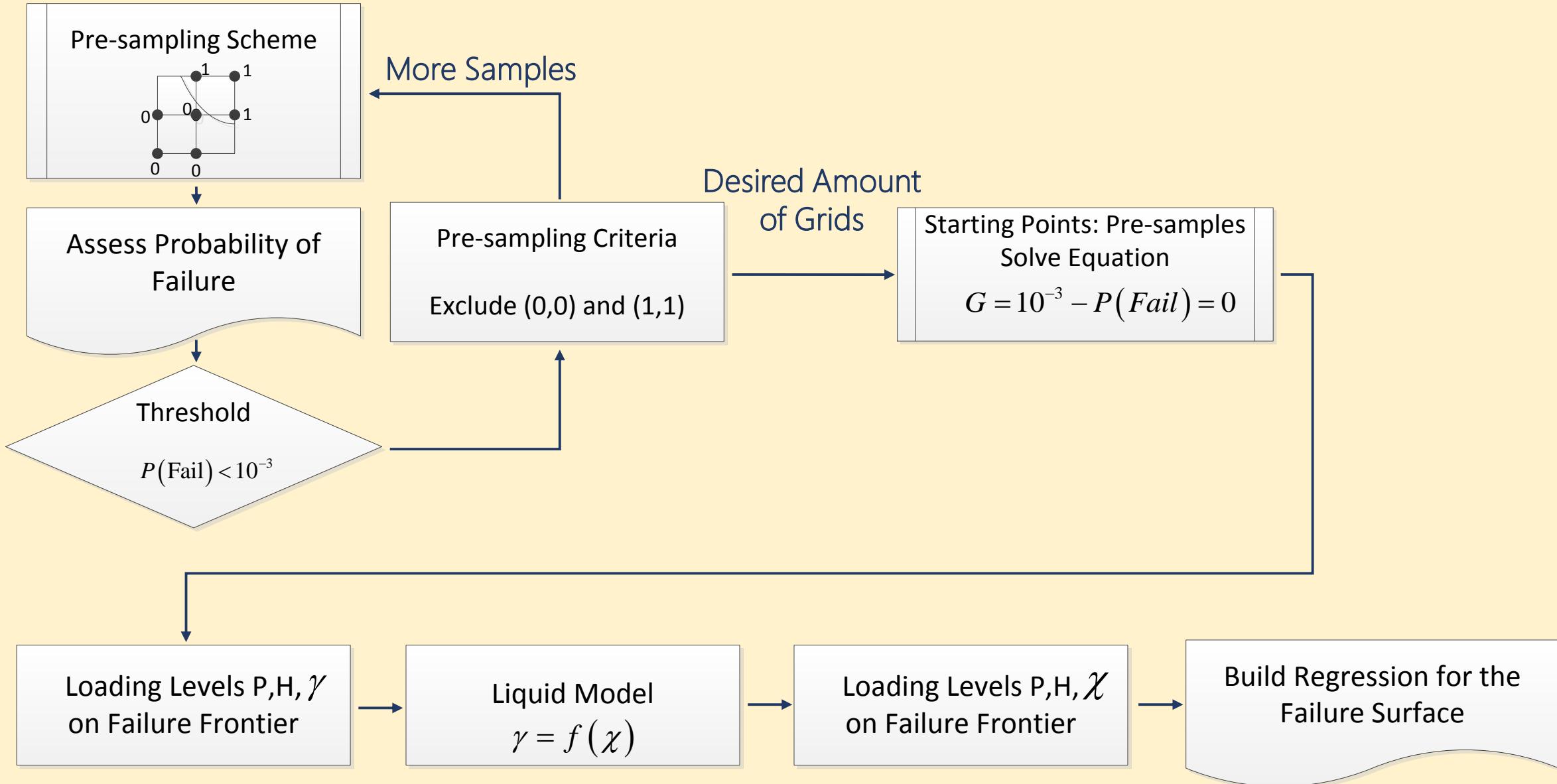
PRE-SAMPLING SCHEME

- The failure frontier
- Highest loading levels
- Lowest loading levels
- Safe
- Fail

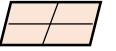
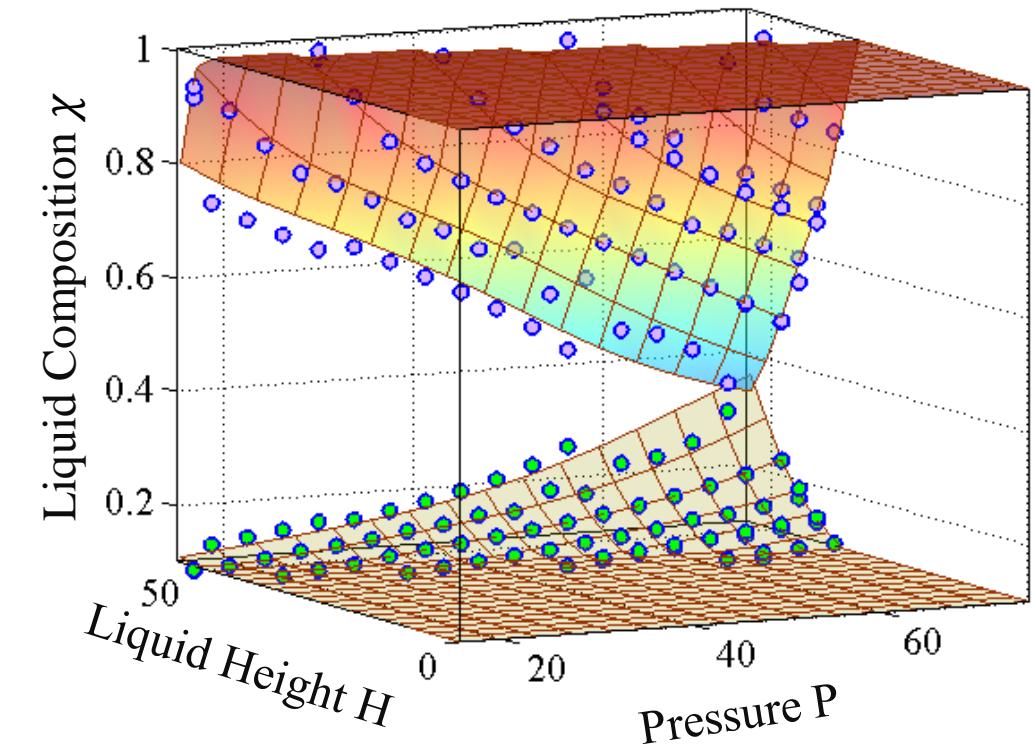
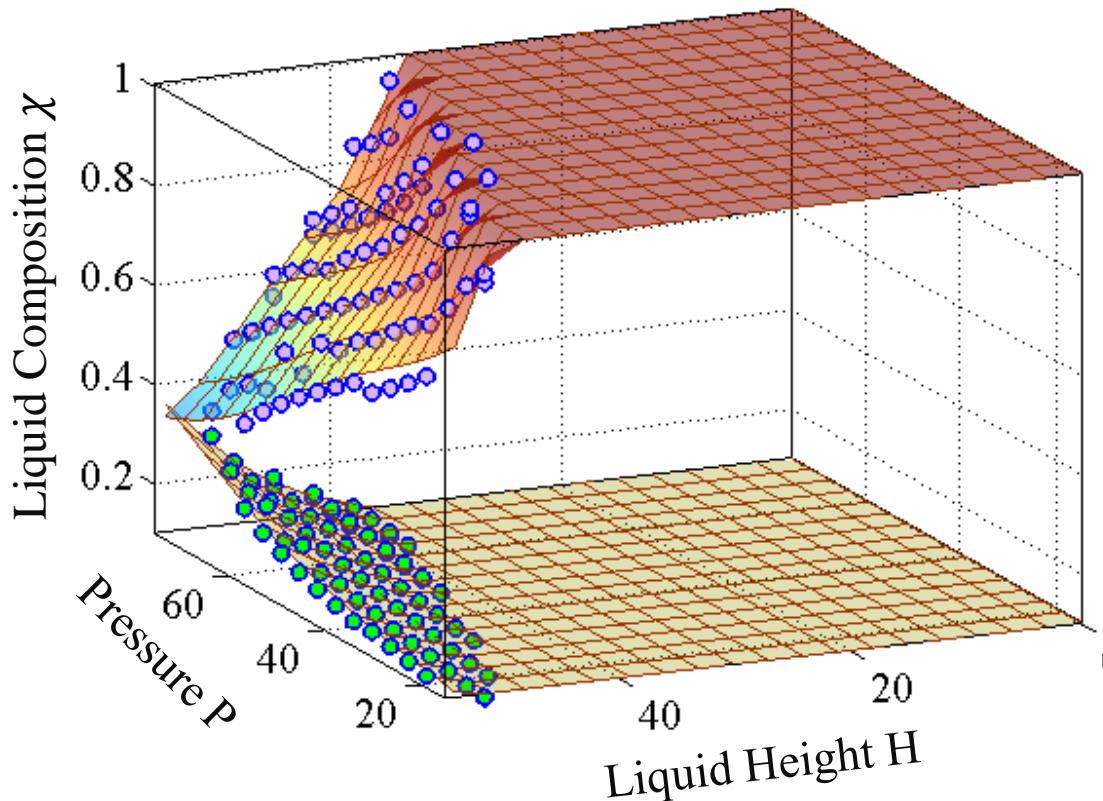


Criterion: Exclude the areas with same Safe (0,0) or Fail (1,1) code.

> Finding the Limit State in Operation Space



> Results

 Failure Frontiers P, H, and χ for Regression

Summary

> Pros & Cons of the Approach

Summary

Challenges	Our Approach	Benefits	Potential Issues
Multiple Sources of Uncertainty	Data Preprocessing	Reduce experimental uncertainty	None
	Sensitivity Analysis	Reduce parameter uncertainty	None
	SRP-Based Model Calibration and Bias Correction	Take into account different sources of uncertainty	Conservative for reliability assessment
No Experimental Data / Expert Opinions	Machine Learning	Allow prediction without experimental data	Need more simulations, introduce new source of uncertainty



Discussions

TIME COMMITMENT

- < 2 months

SIMULATION COSTS

- Calibration: 100 runs
- Model Bias Correction: 200 runs
- Machine Learning: 1,000~10,000 runs

SUGGESTIONS FOR FUTURE PROBLEM

- Provide a few experimental data for validation
- Provide experts' opinions on the failure behavior
- Enable iterative analysis-experiment procedure

POTENTIAL IMPROVEMENTS OF OUR ANALYSIS

- Multi-fidelity uncertainty quantification
- Multi-response model calibration & bias correction (if stress data were available)

Thank You!

Zhen Jiang PhD Candidate
ZhenJiang2015@u.northwestern.edu

Wei Chen Professor
WeiChen@northwestern.edu